DOCUMENT RESUME

ED 124 193 IR 003 586

AUTHOR Christie, Stephen
TITLE New Methods in Information Retrieval.
INSTITUTION British Broadcasting Corp., London (England).
PUB DATE May 76
NOTE 6p.; Paper presented at the National Information Retrieval Colloquium (Philadelphia, Pennsylvania, May 4-6, 1976)

EDRS PRICE MF-$0.83 HC-$1.67 Plus Postage.
DESCRIPTORS Broadcast Television; *Facsimile Communication Systems; *Innovation; *News Media; *Television
IDENTIFIERS British Broadcasting Corporation; *CEEFAX; Great Britain

ABSTRACT

CEEFAX, the British Broadcasting Corporation's news and information service, allows viewers to select from and read 100 pages of reporting written in words and figures on the television screen. A receiver equipped with CEEFAX has extra controls with which to select the service and to dial the page. This optional service requires a built-in decoder to supplement the television receiver. The volume of information shown can be increased in various ways, and subtitles can be superimposed on normal programs. Illustrations and a brief specification of the system are included. (EMH)

****
* Documents acquired by ERIC include many informal unpublished *
* materials not available from other sources. ERIC makes every effort *
* to obtain the best copy available. Nevertheless, items of marginal *
* reproducibility are often encountered and this affects the quality *
* of the microfiche and hardcopy reproductions ERIC makes available *
* via the ERIC Document Reproduction Service (EDRS). EDRS is not *
* responsible for the quality of the original document. Reproductions *
* supplied by EDRS are the best that can be made from the original. *
****

ED124193

New Methods in Information Retrieval

by

Stephen Christie

U S DEPARTMENT OF HEALTH,
EDUCATION & WELFARE
NATIONAL INSTITUTE OF
EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

Paper presented at the 13th Annual National Information Retrieval Colloquium
Philadelphia, Pennsylvania
May 4-6, 1976

IR003586

BBC

Engineering Information

Information Sheet 4008(5) October 1975

New Methods in Information Retrieval

by Stephen Christie

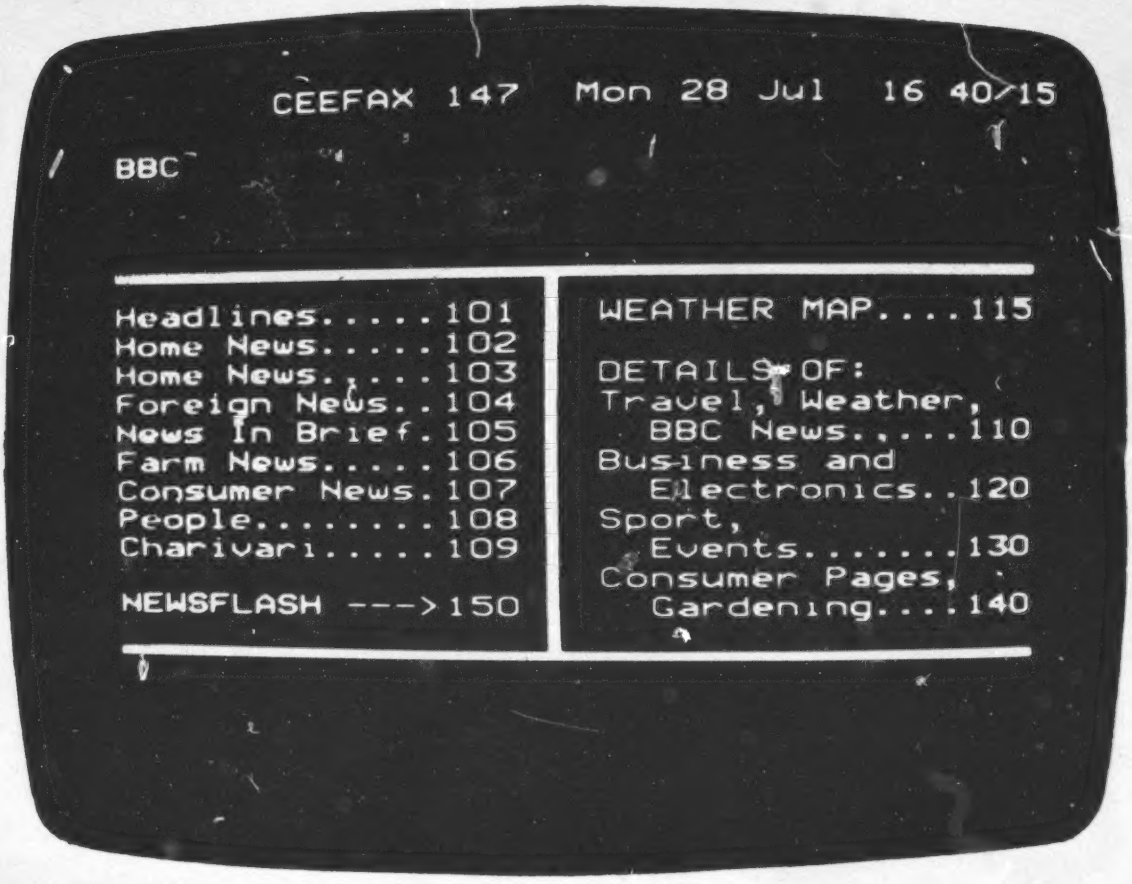

Figure 1

CEEFAX is the BBC's 'dial-a-page' news and information service. It lets the viewer 'see facts' written in words and figures on the television screen. The service is an optional extra and to use it television receivers need a special, built-in decoder. Eventually, a receiver incorporating CEEFAX will cost relatively little more than a standard television set.

A television receiver equipped with a CEEFAX decoder has two extra controls that enable the viewer to switch from the ordinary programme and choose one of the 100, coloured, screen-size pages the CEEFAX magazine can provide. News headlines, sports results, weather, travel, Stock Market prices, consumer affairs–CEEFAX covers all of these and more. All the viewer does is to press a button to select the service then dial a three-figure number for the page required.

Figure 2

The CEEFAX Weathermap: like the other pages, it can have up to six colours along with white.

CEEFAX was first made known in an announcement by the BBC on 23 October 1972 and experimental transmissions began in March 1973. After extensive collaboration with other interested organizations, a United Kingdom unified specification for such an information service was agreed and published*. BBC test transmissions changed to the unified standard in April 1974 and an experimental service carrying many different pages, continually updated by a team of editors and researchers began, with the authorization of the Home Office, on September 23, 1974.

A CEEFAX page

Figures 1 and 2 show what CEEFAX pages look like. Every page can have 24 rows of information with up to 40 characters† in every row. The shape of the characters will depend on receiver design but the system is tailored to a 7 x 5 matrix.

The first row of every page is a special row called the 'page header'. The page header will show the page number, the date and the time to the second. The CEEFAX receiver recognises signals transmitted at the beginning of the header row so that it can tell when a page is beginning and which page it is. When the signals match the orders given by the viewer through his page selector the receiver will proceed to display the header row and the text which follows it.

* **Specification of standards for information transmission by digitally coded signals in the field-blanking interval of 625-line television systems**, published jointly by the British Broadcasting Corporation, the Independent Broadcasting Authority (IBA) and the British Radio Equipment Manufacturers' Association (BREMA), price 50p, post free. CEEFAX is the BBC's name for the information service; it is known as Teletext to BREMA and ORACLE to the IBA.

† except the page header which has only 32 characters.

Transmission

The information is transmitted as a series of coded electronic pulses. Figure 3 shows how the system operates. The pulses are added to two lines in the field blanking period of the television waveform (lines 17 and 18 in one field and 330 and 331 in the next). Each of these television lines carries the coded information for one row of the CEEFAX display. Thus with two lines available for each television field it will take

$$\frac{24 \text{ rows}}{2 \text{ rows/field} \times 50 \text{ field/sec}} = 0.24 \text{ seconds}$$

to transmit a full page of 24 rows. For good legibility and attractive layout, the average page will contain less than 24 rows of text and is therefore transmitted in less than 0·24 seconds. Complete pages are transmitted one after the other and it may take up to 24 seconds after pushing the selector buttons before the required page appears on the screen.

Reception

Like ordinary television, CEEFAX can be disturbed by severe interference but the effect is to produce incorrect characters. However, good CEEFAX reception is possible wherever good television pictures can be received and field tests have shown that CEEFAX works equally well on uhf and vhf transmissions.

Figure 4 shows, very simply, the arrangement of a CEEFAX receiver.

Extending a magazine

The system allows for many refinements. The amount of information that can be shown can be considerably extended by using the pages in the following ways.

Type A pages are single pages of information that can be rapidly updated. News headlines are shown in this way.

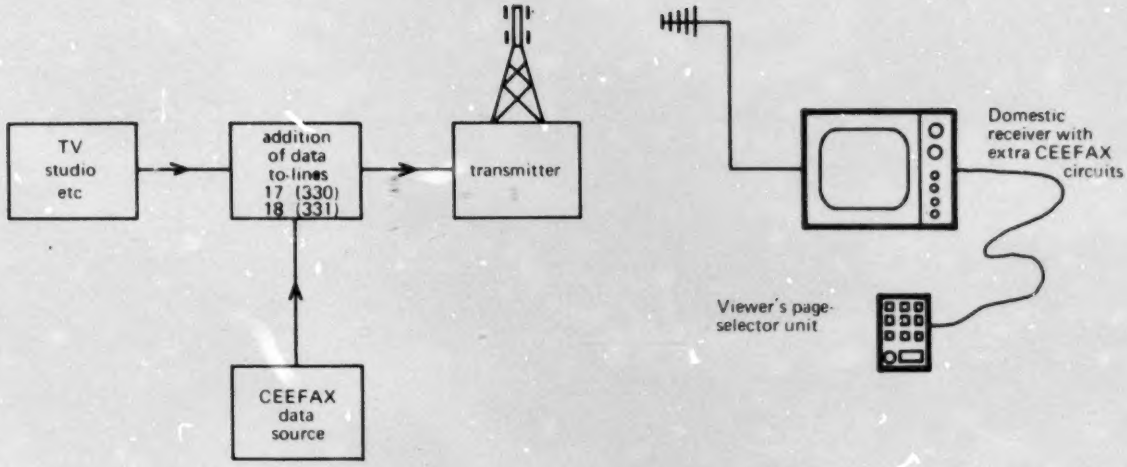

Figure 3: CEEFAX basic arrangement

Type B pages can be described as 'rotating' pages. A complete page of information is transmitted and left unchanged for a reasonable reading period, perhaps one minute. Then a new text, following on from the first, automatically takes its place. In this way long lists, for example of football results, take up the minimum space in the system.

Type C pages are used to carry information which does not need frequent updating and might require to be transmitted only once each day. At the rate of one a minute there can be more than 1,000 'editions' of a particular page in a day, each identified by its own 'address' which, for convenience, is also the time at which it is transmitted. For example, a page addressed 1925 is transmitted at 1925 clock-time. The type C page is preselected by the viewer so that his receiver recognises and stores the required page, to be viewed later at some convenient time. It is conceivable that electronic stores capable of storing many such pages, throughout the day, will become available at an economic price.

CEEFAX can also be used to superimpose subtitles on the normal programmes. This could be useful for deaf viewers or could give an alternative language version of the programme sound. In a similar way the viewer can choose to have newsflashes automatically displayed as soon as they are transmitted.

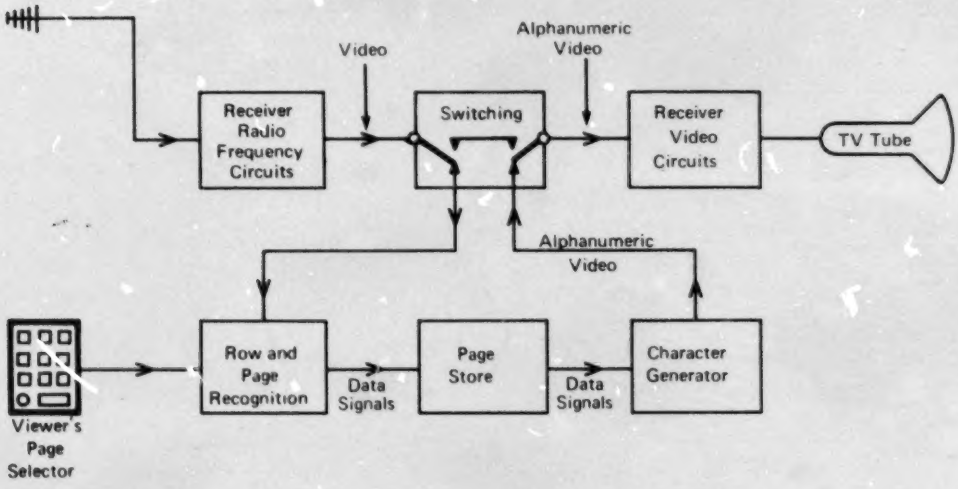

Figure 4: CEEFAX receiver arrangement

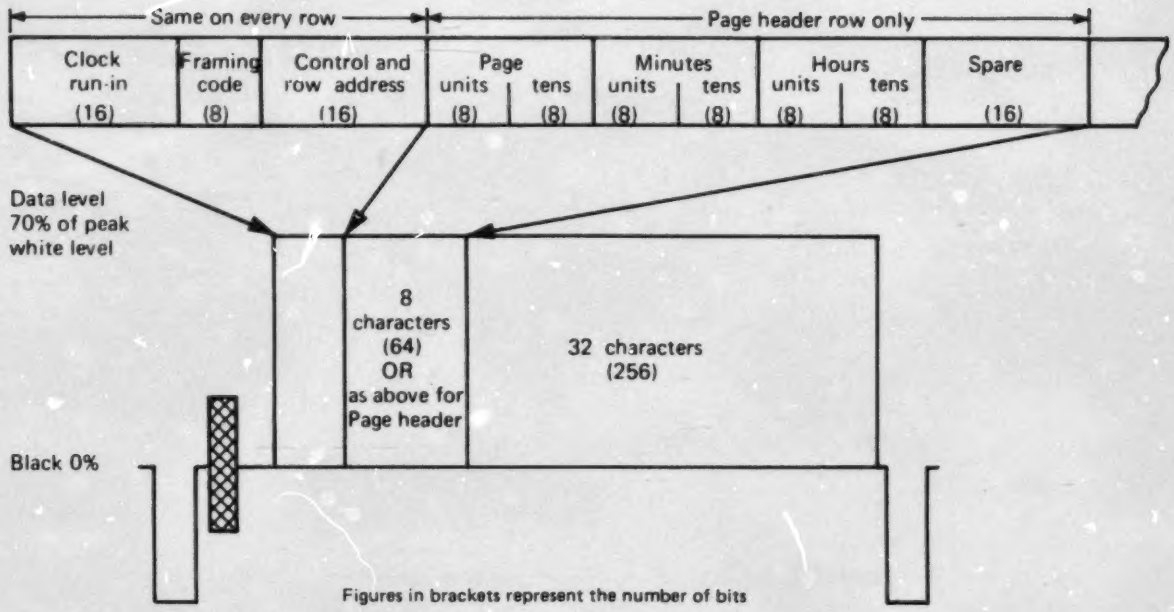

Figure 5: The organization of CEEFAX data on lines 17, 18, 330 and 331

The CEEFAX waveform

1. **Pulses**
   Shape–raised cosine
   Amplitude–approzimately 70% of the black to white excursion
   Clock rate–6·9375 MHz
   Logic–positive: a data pulse corresponds to '1'
   no pulse corresponds to '0'
   System coding–NRZ (non-return-to-zero): a simple 'on/off' binary code

2. **Pulse Grouping**
   Active line time is divided as follows:-
   (see also Fig. 4)

   2.1 *Clock and identification pulses*
   Clock run-in– 16 bits are allocated to a simple sequence of pulses that brings the receiver 'clock' into step with the transmitted data.
   Framing-code– the framing code is to allow the receiver to achieve word synchronism with the data which follows.

   2.2 *Data pulses*
   Information carrying pulses fall into two groups; the addresses, which tell the receiver the number of the row and page being transmitted, and the characters. Each group uses different transmission codes.

   2.2.1
   Addresses of pages and rows are transmitted in binary coded decimal (BCD). This simplifies the receiver circuits needed to detect a match between the receiver's page selector and the number of the page being transmitted. A Hamming system is used to protect each BCD digit which then occupies 8 bits. The Hamming technique enables single errors to be corrected. Addresses transmitted in this way are:-

   Control and row address– enables the receiver to identify the row being transmitted so that each row of information is displayed in the correct location on the screen. With an address for each row, blank rows need not be transmitted.
   The row address– also serves to indicate the beginning of a page.
   Page address– The page address occupies part of the page header row. The address (in BCD) indicates the page number and the time of transmission (to the minute). These factors initiate storage and display of the selected CEEFAX page.

   2.2.2
   Characters are transmitted in a seven-bit code which allows for a wide range of alpha-numeric characters. An eighth parity bit is added for error protection. Odd parity is used.
   One television line can carry a complete row of 40 characters but when the row is the page header only 32 characters can be accommodated because the page address (see above) takes up 64 bits (equivalent to 8 characters).

3. **Pages and magazines**
   The specification of the system allows for eight magazines with up to 100 pages in each magazine.